Benefits and Pitfalls of GRACE Data Assimilation: a Case Study of Terrestrial Water Storage Depletion in India

Manuela Girotto[1,2], Gabriëlle J. M. De Lannoy[3], Rolf H. Reichle[1], Matthew Rodell[4], Clara Draper[5,6], Soumendra N. Bhanja[7], Abhijit Mukherjee[7,8]

[1]Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, MD, USA.

[2]GESTAR, Universities Space Research Association, Columbia, MD 21044, USA.

[3]Department of Earth and Environmental Sciences, KU Leuven, Belgium.

[4]Hydrological Science Lab, NASA Goddard Space Flight Center, Greenbelt, MD, USA.

[5]Physical Sciences Division, NOAA/Earth System Research Laboratory, Boulder, Colorado.

[6]Cooperative Institute for Research in the Environmental Sciences, University of Colorado Boulder, Boulder, Colorado.

[7]Department of Geology and Geophysics, Indian Institute of Technology, Kharagpur, India.

[8]School of Environmental Science and Engineering, Indian Institute of Technology Kharagpur, West Bengala, India.

Key Points:

- GRACE observations of terrestrial water storage (TWS) in northwest India show trends likely associated with groundwater extraction.
- Land models in global assimilation systems do not usually represent anthropogenic processes such as groundwater extraction and irrigation.
- Assimilation of GRACE observations introduces realistic trends in TWS and groundwater along with an erroneous trend in evapotranspiration.

Corresponding author: Manuela Girotto, manuela.girotto@nasa.gov

Abstract

This study investigates some of the benefits and drawbacks of assimilating Terrestrial Water Storage (TWS) observations from the Gravity Recovery and Climate Experiment (GRACE) into a land surface model over India. GRACE observes TWS depletion associated with anthropogenic groundwater extraction in northwest India. The model, however, does not represent anthropogenic groundwater withdrawals and is not skillful in reproducing the interannual variability of groundwater. Assimilation of GRACE TWS introduces long-term trends and improves the interannual variability in groundwater. But the assimilation also introduces a negative trend in simulated evapotranspiration whereas in reality evapotranspiration is likely enhanced by irrigation, which is also unmodeled. Moreover, in situ measurements of shallow groundwater show no trend, suggesting that the trends are erroneously introduced by the assimilation into the modeled shallow groundwater, when in reality the groundwater is depleted in deeper aquifers. The results emphasize the importance of representing anthropogenic processes in land surface modeling and data assimilation systems.

1 Introduction and Background

India is the world's largest user of groundwater resources [*Aeschbach-Hertig and Gleeson*, 2012], and irrigation accounts for more than 85% of its groundwater withdrawals [*FAO*, 2013]. The current rate of groundwater consumption is unsustainable and may eventually increase poverty and food insecurity in rural India [*Zaveri et al.*, 2016]. Monitoring these risks is essential in this era of rapid socio-economic growth and climate change. This will require an improved understanding of the factors that affect groundwater and of the relationship between groundwater and other components of the water cycle such as soil moisture, vegetation, precipitation and evapotranspiration.

Global assessment of groundwater depletion and variations has been facilitated by observations from the Gravity Recovery and Climate Experiment (GRACE) satellite mission [*Tapley et al.*, 2004]. GRACE provides monthly, vertically-integrated estimates of terrestrial water storage (TWS) anomalies (departures from the long-term mean), at coarse spatial scales (~300 km). TWS comprises groundwater, soil water, surface water, snow, and ice. GRACE observations have been used to estimate groundwater depletion rates around the world [*Famiglietti and Rodell*, 2013]. In particular, *Rodell et al.* [2009]; *Tiwari et al.* [2009]; *Shamsudduha et al.* [2012]; *Panda and Wahr* [2016], studied groundwater depletion in India based on GRACE TWS observations. In these studies, groundwater was isolated from the observed (GRACE) TWS by subtracting independent estimates of surface water and offline (land-only) model estimates of soil water, snow, and ice. The effects of groundwater depletion and irrigation on soil moisture and evapotranspiration were not assessed.

Assimilation of GRACE observations into a land surface model permits investigation of the impacts of groundwater depletion on other water storage compartments and the fluxes between them. It also enables spatial, vertical, and temporal disaggregation of the TWS components, including groundwater, surface and root zone soil moisture and snow [*Zaitchik et al.*, 2008], while preserving the internal consistency of the modeled storages and fluxes and taking into account uncertainties due to model and observational errors. Model uncertainty is caused by errors in surface meteorological forcing, model parameters, and model structural errors. Some of the uncertainty is related to unmodeled processes, most notably human impacts such as pumping from aquifers, irrigation, or water management [*Ozdogan et al.*, 2010]. Further, it is common to rescale the observations prior to data assimilation in order to address model and observation biases (e.g., *Reichle*

and Koster [2004]). However, such rescaling may discard important signals in the observations [*Kumar et al.*, 2015]. Thus, a remaining challenge in data assimilation is to isolate errors caused by unmodeled processes so that the true observational features are not excluded during data assimilation [*Kumar et al.*, 2015].

In this study, we investigate the extent to which GRACE data assimilation can overcome modeling errors, including errors that arise from the lack of representation of groundwater extraction and irrigation. Simulated TWS, groundwater, and evapotranspiration are evaluated over India, where the assimilated GRACE TWS observations contain trends due to the ongoing groundwater depletion, an anthropogenic and unmodeled process. Benefits and drawbacks of the assimilation scheme are evaluated in terms of its ability to improve simulated seasonal and interannual variability and trends.

2 Methods and Data

2.1 Model and Forcings

Consistent with *Girotto et al.* [2016], this work uses the Catchment land surface model (CLSM, *Koster et al.* [2000]) and Modern Era Retrospective Analysis for Research Application (MERRA) meteorological forcing data [*Rienecker et al.*, 2011]. CLSM is one of the few widely used land surface models that includes a basic representation of shallow (unconfined) groundwater storage variations (*Koster et al.* [2000]; their Figure 2). However, it does not simulate deeper multilayer aquifers or dynamic surface water hydrology (e.g., lakes and rivers). The study domain encompasses India and Bangladesh and covers January 2003 to December 2015. The simulations are performed on a 36-km Equal Area Scalable Earth (version 2) grid [*Brodzik et al.*, 2012].

2.2 GRACE Terrestrial Water Storage Observations

The Level-3, monthly, 1°x1° gridded, spherical harmonic based GRACE TWS product available from the Jet Propulsion Laboratory (http://grace.jpl.nasa.gov) is used. The data are a truncated and smoothed [*Landerer and Swenson*, 2012] version of the RL05 solution from the Center for Space Research at the University of Texas. Prior to data assimilation, we rescale the GRACE TWS observations to match the long-term mean and standard deviation of the model [*Girotto et al.*, 2016]. This does not imply that the climatology of the model is more correct than that of the observations; it is done to remove the long-term systematic bias in the mean and variance between the model and the observations while preserving trends and seasonal-to-interannual variations in the rescaled observations.

2.3 Data Assimilation

The assimilation system is fully described in *Girotto et al.* [2016]. Here, only the key points and differences are noted. A 3D ensemble Kalman Filter (EnKF) is used, where the “3D" notation refers to the fact that the filter distributes information horizontally as well as vertically [*Reichle and Koster*, 2003; *De Lannoy et al.*, 2010]. The assimilation method is similar to an ensemble smoother approach, i.e., it is a “two-step" scheme in which the land model integration is performed twice over the course of the same month: first to collect monthly TWS observation-minus-forecast differences (i.e., innovations), and

a second time to update that month's simulated TWS using increments computed from the observation-minus-forecast residuals obtained in the first integration. The observation predictions are computed by spatially aggregating the monthly TWS estimates from the 36-km model grid using a Gaussian smoothing average function with a 300-km half-width distance (to match the resolution of the GRACE TWS observations; Section 2.2). The ensemble forecast perturbation parameters used here match those reported in *Girotto et al.* [2016] except that we doubled the standard deviation associated with the uncertainty in the "catdef" model prognostic variable (Table S1). This was done because the innovation statistics [*Desroziers et al.*, 2005] indicated that the data assimilation approach required increased model uncertainties (not shown).

2.4 Groundwater in Situ Measurements

The Central Ground Water Board of India measures groundwater levels four times a year during January, April/May, August and November [*CGWB*, 2014]. The data used in this work cover the period from January 2005 to December 2013. Groundwater levels are measured using piezometers in non-pumping wells that are typically located in the shallowest (water table) aquifer and thus represent unconfined or perched aquifers, but not deeper aquifers. Consequently, these measurements are not directly representative of deep aquifers from which groundwater may be extracted, but the data are informative about the human-induced shallow water recharge by irrigation. The data represent equivalent heights of water (i.e., the product of water elevation and specific yield) as described in *Bhanja et al.* [2016]. The data have been quality controlled for temporal continuity and outliers. We aggregated the in situ groundwater measurements from the 3297 well locations to the 36-km model grid, resulting in groundwater validation measurements for 1452 grid cells (out of 2899) within the simulation domain (**Figure 1d**). This abundance of in situ measurement locations is unprecedented for GRACE assimilation studies.

2.5 Trend Analysis and Evaluation Metrics

A modified version of the nonparametric Mann-Kendall test was used to identify the statistical significance of trends in observed and simulated TWS, groundwater, and evapotranspiration, taking into account the temporal autocorrelation in the time series [*Hamed and Ramachandra Rao*, 1998]. The trend magnitude is computed as the median of the slopes calculated from consecutive pairs of sample points [*Sen*, 1968].

Simulated TWS and groundwater are evaluated in terms of time series correlation (R) and anomaly correlation ($anomR$) with observations, and their 95% confidence intervals. The $anomR$ values are calculated after removing both the long-term trends and the mean seasonal cycle from the time series, where the seasonal cycle is calculated as the multi-year average for each calendar month. That is, the R metric is sensitive to trends as well as the seasonal and interannual variability, whereas the $anomR$ metric is sensitive only to the interannual variability. Spatially averaged metrics are computed using a clustering algorithm [*Girotto et al.*, 2016].

3 Results and Discussion

3.1 Trends in TWS and Groundwater

GRACE TWS observations suggest that a significant negative trend exists in northwest India with a maximum rate of -1.7 cm/year near Delhi, a region with intense irrigation (compare trends in **Figure 1a** with areas equipped for irrigation in **Figure 2**). This is consistent with earlier studies [*Rodell et al.*, 2009; *Tiwari et al.*, 2009; *Chen et al.*, 2014], which attributed the trend to groundwater extraction for irrigating crops. A negative trend in TWS (-0.7 cm/year) in the state of Tamil Nadu in southern India (**Figure 1a**) is also ascribed to irrigated agriculture (*Chinnasamy and Agoramoorthy* [2015]). TWS has increased during the study period in west-central India (Maharashtra, Gujarat, and Madhya Pradesh; **Figure 1a**). This region relies more heavily on surface water reservoirs than on groundwater to meet its freshwater needs [*Soni and Syed*, 2015]. The positive trend reflects both a recent increase in precipitation and the filling of reservoirs [*Tiwari et al.*, 2009].

There are no consistent patterns of shallow groundwater trends seen in the in situ data, except in the region of Tamil Nadu (southern India, **Figure 1d**), where a weak negative trend is also present in the TWS observations (**Figure 1a**). On average, trends in the in situ groundwater measurements are mixed to positive, which is in disagreement with GRACE indicating larger areas with a stronger decrease in TWS than increase.

This discrepancy can likely be attributed to differences in the exact quantities observed by GRACE and the situ measurements. Groundwater pumping for irrigation mainly depletes water from the deep aquifers into which most agricultural wells are installed. GRACE cannot distinguish shallow from deep groundwater or other TWS components and lumps them all together as a single quantity. Hence the intense depletion of deep aquifers in northern India dominates the GRACE signal in that region. The in situ groundwater measurements, on the other hand, sample only shallow groundwater (Section 2.4). Moreover, rain and irrigation drainage rapidly percolate to the water table or flow directly into the open wells [*Panda and Wahr*, 2016]. As a result, the in situ measurements do not reflect the long-term changes occurring in the deep aquifers but are useful for evaluating short-term processes (i.e., meteorologically-driven or irrigation enhanced-recharge in shallow aquifers).

The model-only simulation also does not replicate the negative TWS trend in northwest India (**Figure 1b**). By construction, trends are visible in the assimilation case (**Fig-**

ure 1c), consistent with those in the assimilated GRACE TWS observations (**Figure 1a**). For example, the depletion rate in Delhi is -0.75 cm/year in the assimilation case, which is about half of the maximum rate of change in the observed TWS (-1.7 cm/year). Thus, the assimilated result is a compromise between the absence of a trend in the modeled TWS and the GRACE-observed TWS trend.

Likewise, there are no significant trends in the model-only groundwater estimates (**Figure 1e**). GRACE TWS assimilation introduces patterns of groundwater trends (**Figure 1f**) that are comparable to those seen in TWS (**Figure 1c**). For lack of deep aquifers in the Catchment model, the assimilation (perhaps erroneously) introduces the trends in the shallow groundwater, and also (correctly, as will be shown later) updates the groundwater simulations for seasonal and short-term errors. The trend patterns in the assimilation, however, are different from those of the in situ (shallow) groundwater measurements (**Figure 1d**). While there is some agreement in Tamil Nadu (negative trends) and in Madhya Pradesh and Andhara Pradesh (positive trends), no trend is present in the in situ groundwater measurements in northwest India (**Figure 1d**), where the assimilation results indicate strong negative trends (**Figure 1f**).

Figure 3 illustrates, for the location in northwest India with the strongest TWS trend, the assimilated GRACE TWS observations along with groundwater estimates from the independent in situ measurements, the model-only, and the assimilation estimates. All time series show a similar amplitude and phase of the seasonal cycle (**Figure 3a**). GRACE indicates a strong negative TWS trend, which is not simulated by the model and is also not observed in the shallow groundwater measurements. The assimilation corrects the overly dry modeled groundwater estimates during 2003-2005, but it fails to adjust the overly wet model estimates towards the very dry TWS observations during 2010-2016. The latter is a consequence of a lower limit in modeled TWS, which is determined by the prescribed depth-to-bedrock [*Houborg et al.*, 2012; *Li et al.*, 2012].

Anomalies in GRACE TWS and in situ groundwater measurements (after removing secular trends and the seasonal cycle) indicate dry conditions (negative anomalies) during 2007, 2009 and 2010, while the model-only experiment indicates near-normal conditions in those years (**Figure 3b**). GRACE data assimilation induces negative TWS and groundwater anomalies in those years, thereby improving the agreement between simulated and observed groundwater. Likewise, the GRACE-observed wet period during winter 2003-2004 is underestimated by the model and corrected by the assimilation (**Figure 3b**).

3.2 Trends in Evapotranspiration Fluxes

We evaluated trends in additional water budget components. For example, an analysis of soil moisture yields similar conclusions to those found for the model-only and assimilation groundwater results (Section 3.1). Important additional insights are gained by investigating evapotranspiration. While there are no significant trends in the model-only evapotranspiration (**Figure 1h**), significant trends are seen in the assimilated evapotranspiration (**Figure 1i**) which mimic the TWS trends (**Figure 1c**). Trend patterns based on independent evapotranspiration datasets, e.g., Jung et al. [2009] (**Figure 1g**) contradict the assimilation results. The negative evapotranspiration trends in northern India in **Figure 1i** are a direct consequence of the water deficit induced by the assimilation of the GRACE-observed negative TWS anomalies. In reality, irrigation likely sustains root-zone moisture (as indirectly suggested by the shallow groundwater measurements) and allows evapotranspiration to continue at a steady (or even increased) rate. While the assimilation of TWS for areas with a natural water budget should, in theory, improve the accuracy of evapotranspiration variations (provided natural processes are adequately represented in the model), the inability of the model to simulate groundwater-supported irrigation in this case caused a degradation of simulated evapotranspiration when TWS was assimilated.

3.3 Correlation Metrics

In this section we report correlation (*R*, *anomR*) metrics of model-only and assimilation results versus the assimilated GRACE TWS observations and versus the independent in situ groundwater measurements. For reference, the supplemental material provides maps of the long-term precipitation and TWS climatologies (**Figure S1**). We refer to wet and dry areas where the annual mean precipitation is more or less, respectively, than the average over India (**Figure S1a**).

3.3.1 Terrestrial Water Storage

In general, higher *R* values between modeled and GRACE TWS are found in the wetter parts of India (compare **Figure 4a** with **Figure S1a**), where the seasonal and interannual variability is stronger and where weaker or no human-induced trends from groundwater pumping and irrigation are expected. An exception is the wet region of southern India, where the seasonal cycle of precipitation is bimodal, resulting in higher errors in the modeled TWS time series, and thus lower *R*. Lower *R* values are generally found in the drier regions, where (i) the interannual and seasonal variability of both the GRACE and modeled TWS are lowest, as suggested by their long-term standard deviation (**Figure S1b-c**), or where (ii) trends and interannual variability are affected by anthropogenic processes which are not modeled, but reported by the GRACE observations (**Figure 1a**). By design, the GRACE data assimilation increases the *R* between the simulations and GRACE to a domain-average of 0.96, compared to 0.83 prior to assimilation, with the largest increase in *R* in drier regions (compare **Figure 4b** with **Figure S1a**), where the model fails to represent human-induced trends.

The highest TWS *anomR* values are in the central wetter regions of India (e.g., Maharashtra, Madhya Pradesh, Orissa, West Bengal; compare **Figure 4c** with **Figure S1a**). The lowest *anomR* values are in the northwest (e.g., Punjab, Haryana, New Delhi) and in the south (Tamil Nadu). Low *anomR* values indicate poor model interannual variability representation, possibly due to the lack of irrigation modeling. By design, the assimilation strongly increases the *anomR* over the entire region (**Figure 4d**) to a domain average value of 0.90, versus 0.51 prior to assimilation. The largest increases are in the northwest and in Tamil Nadu, where anthropogenic processes affect the hydrologic interannual variability. The assimilation only marginally increases the *anomR* in the wet regions of the domain, where irrigation is less likely to regulate the water budget (**Figure 2**).

3.3.2 Groundwater

The domain-average (with 95% confidence interval) *R* between model-only groundwater estimates and independent, in situ groundwater measurements equals R=0.51±0.05 (**Figure 4e**). The lowest correlations are in the north (i.e., Rajasthan, Haryana, Delhi), south (i.e., Tamil Nadu), and east (i.e., Assam) of India. Similar to the TWS evaluation (**Figure 4a**), model performances are higher in the wet regions (compare **Figure 4e** with **Figure S1a**), where the seasonal and interannual variability is less affected by antropogenic interventions and where the model can reproduce the natural variability.

GRACE TWS assimilation improves groundwater *R* in a majority (73%) of the in situ locations, such as Tamil Nadu (**Figure 4f**), but it degrades groundwater fidelity in some locations (e.g., northwest Orissa, north Rajasthan). Overall, the domain-average improvement in *R* is 0.05 (not statistically significant), resulting in R=0.56±0.05 for the assimilation estimates. Improvements may be attributed to better representation of seasonal and interannual variability. This positive increase in the statistics corroborates the findings of *Girotto et al.* [2016], who demonstrated that the downscaling of vertically integrated and spatially coarse-scale GRACE TWS generally improves the simulation of groundwater at finer scales.

The *anomR* between model-only groundwater and in situ measurements is consistently very low, with a domain average *anomR*=0.13±0.06 (**Figure 4g**). Higher values (*anomR* >0.4) are found in the states of Gujarat and Maharashtra, where irrigation intensity is low (**Figure 2**). The interannual variability of the in situ groundwater measurements is, in general, not well replicated by the model, possibly because the model does not simulate irrigation. The strongest improvements in simulated groundwater induced by GRACE data assimilation are in north-central India (Madhya Pradesh, Bihar Jharkhand) and in south-central India (Tamil Nadu, Kernataka; **Figure 4h**). Skill is degraded at some locations scattered throughout the country, including a cluster in the western states of Assam, Orissa and Gujarat (**Figure 4h**). Nonetheless, on average the skill of the assimilation estimates is improved to *anomR*=0.23±0.06. These improvements imply that GRACE data assimilation can enhance the interannual variability of simulated groundwater in the presence of anthropogenic processes. However, despite the relatively large *anomR* increase of 0.10, the improvement is still not statistically significant, because of the low *anomR* values and the limited number of monthly sample points for validation. In any case, the very low

skill highlights the urgent need to improve the model representation of deep groundwater and of pumping and irrigation processes.

4 Conclusions

Anthropogenic processes are often not included in global land surface modeling systems, but regional patterns in groundwater extraction and irrigation over India are observed by the GRACE satellite mission. This paper investigates the extent to which GRACE data assimilation can correct (or not) for errors due to missing model processes.

The GRACE observations show strong negative TWS trends in northwest India, and weaker negative trends in Tamil Nadu. These trends are caused by the depletion of groundwater for irrigation purposes (e.g., *Rodell et al.* [2009]). In situ shallow groundwater measurements show clear trends only in southern India (Tamil Nadu). In general, the in situ groundwater trends are not regionally uniform and are inconsistent with the GRACE TWS observations. We attribute this difference to the fact that groundwater used for irrigation is extracted primarily from deep aquifers, which are observed by GRACE, but not by the (shallow) in situ groundwater measurements.

The model-only simulation does not include groundwater extraction and therefore does not reproduce the significant GRACE-observed TWS trends in India. The assimilation of GRACE TWS observations introduces trends in the modeled TWS and groundwater. But the model does not simulate deeper aquifers, and, consequently, the assimilation assigns the water storage updates to the model's shallow groundwater compartment. The result is a crude but not entirely inaccurate accounting of vertically integrated groundwater storage variations. One unintended consequence, however, is that the GRACE assimilation unrealistically reduces evapotranspiration, because the model also does not simulate irrigation.

The highest correlations (R) and anomaly correlations ($anomR$) between the model-only and GRACE-observed TWS are in the wetter parts of India, where the seasonal and interannual variability is more dominated by natural, rather than anthropogenic, processes. By construction, GRACE data assimilation leads to better correlations with GRACE TWS observations.

We further evaluated the results in terms of the R and $anomR$ values versus the (shallow) in situ groundwater measurement, which sample about half of the domain. Both the model-only and assimilation estimates have very low $anomR$ versus the groundwater observations. We attribute this to: (1) the lack of simulation by the model of irrigation and irrigation return flows, (2) the fact that the in situ measurements observe only shallow groundwater and thus are not representative of the total column groundwater changes

observed by GRACE, and (3) the limitation in the dynamic range of the modeled groundwater that is imposed by its depth-to-bedrock parameter.

Despite the model's shortcomings, GRACE data assimilation produces improvements (not statistically significant) in groundwater R and $anomR$ even in areas that are strongly affected by anthropogenic and unmodeled processes. Finally, these results should motivate the land surface modeling and data assimilation community to better represent anthropogenic impacts on the water cycle by adding the relevant processes into the model, including the simulation of irrigation, groundwater extraction, and deep subsurface water storage variations.

Acknowledgments

This study was supported by the NASA Terrestrial Hydrology and GRACE & GRACE Follow-On Science Team programs. The GRACE observations used here are from http://grace.jpl.nasa.gov, which is supported by the NASA MEaSUREs Program. This manuscript uses open-source data of the Central Groundwater Board (CGWB), Ministry of Water Resources, River Development and Ganga Rejuvenation, Government of India. S. Bhanja was supported by the SPM fellowship, CSIR (India). A. Mukherjee was supported by the Ministry of Human Resource Development (MHRD) project AGI (IIT/SRIC/GG & CSE/AGI/2013-14/201), application of artificial intelligence in groundwater storage estimation of Indian subcontinent. Computational resources were provided by the NASA High-End Computing Program through the NASA Center for Climate Simulation at the Goddard Space Flight Center.

References

Aeschbach-Hertig, W., and T. Gleeson (2012), Regional strategies for the accelerating global problem of groundwater depletion, *Nature Geoscience*, *5*(12), 853–861.

Bhanja, S. N., A. Mukherjee, D. Saha, I. Velicogna, and J. S. Famiglietti (2016), Validation of GRACE based groundwater storage anomaly using in-situ groundwater level measurements in India, *Journal of Hydrology*.

Brodzik, M. J., B. Billingsley, T. Haran, B. Raup, and M. H. Savoie (2012), Ease-grid 2.0: Incremental but significant improvements for earth-gridded data sets, *ISPRS International Journal of Geo-Information*, *1*(1), 32–45.

CGWB (2014), Groundwater year book, *India:Ministry of Water Resources*.

Chen, J., J. Li, Z. Zhang, and S. Ni (2014), Long-term groundwater variations in northwest india from satellite gravity measurements, *Global and Planetary Change*, *116*, 130–138.

Chinnasamy, P., and G. Agoramoorthy (2015), Groundwater storage and depletion trends in tamil nadu state, india, *Water Resources Management*, *29*(7), 2139–2152.

De Lannoy, G. J., R. H. Reichle, P. R. Houser, K. R. Arsenault, N. E. Verhoest, and V. R. Pauwels (2010), Satellite-scale snow water equivalent assimilation into a high-resolution land surface model, *Journal of Hydrometeorology*, *11*(2), 352–369.

Desroziers, G., L. Berre, B. Chapnik, and P. Poli (2005), Diagnosis of observation, background and analysis-error statistics in observation space, *Quarterly Journal of the Royal Meteorological Society*, *131*(613), 3385–3396.

Famiglietti, J. S., and M. Rodell (2013), Water in the balance, *Science*, *340*(6138), 1300–1301.

FAO (2013), Yearbook, world food and agriculture, *Food and Agriculture Organization of the United Nations, Rome*.

Girotto, M., G. J. De Lannoy, R. H. Reichle, and M. Rodell (2016), Assimilation of gridded terrestrial water storage observations from GRACE into a land surface model, *Water Resources Research*.

Hamed, K. H., and A. Ramachandra Rao (1998), A modified Mann-Kendall trend test for autocorrelated data, *Journal of Hydrology*, *204*(1), 182–196.

Houborg, R., M. Rodell, B. Li, R. Reichle, and B. F. Zaitchik (2012), Drought indicators based on model-assimilated gravity recovery and climate experiment (GRACE) terres-

trial water storage observations, *Water Resources Research*, *48*(7).

Jung, M., M. Reichstein, and A. Bondeau (2009), Towards global empirical upscaling of fluxnet eddy covariance observations: validation of a model tree ensemble approach using a biosphere model, *Biogeosciences*, *6*(10), 2001–2013.

Koster, R. D., M. J. Suarez, A. Ducharne, M. Stieglitz, and P. Kumar (2000), A catchment-based approach to modeling land surface processes in a general circulation model: 1. model structure, *Journal of Geophysical Research: Atmospheres (1984–2012)*, *105*(D20), 24,809–24,822.

Kumar, S., C. Peters-Lidard, J. Santanello, R. Reichle, C. Draper, R. Koster, G. Nearing, and M. Jasinski (2015), Evaluating the utility of satellite soil moisture retrievals over irrigated areas and the ability of land data assimilation methods to correct for unmodeled processes, *Hydrology and Earth System Sciences*, *19*(11), 4463.

Landerer, F., and S. Swenson (2012), Accuracy of scaled GRACE terrestrial water storage estimates, *Water Resources Research*, *48*(4).

Li, B., M. Rodell, B. F. Zaitchik, R. H. Reichle, R. D. Koster, and T. M. van Dam (2012), Assimilation of GRACE terrestrial water storage into a land surface model: Evaluation and potential value for drought monitoring in western and central europe, *Journal of Hydrology*, *446*, 103–115.

Ozdogan, M., M. Rodell, H. K. Beaudoing, and D. L. Toll (2010), Simulating the effects of irrigation over the united states in a land surface model based on satellite-derived agricultural data, *Journal of Hydrometeorology*, *11*(1), 171–184.

Panda, D. K., and J. Wahr (2016), Spatiotemporal evolution of water storage changes in india from the updated grace-derived gravity records, *Water Resources Research*, *52*(1), 135–149.

Reichle, R. H., and R. D. Koster (2003), Assessing the impact of horizontal error correlations in background fields on soil moisture estimation, *Journal of Hydrometeorology*, *4*(6), 1229–1242.

Reichle, R. H., and R. D. Koster (2004), Bias reduction in short records of satellite soil moisture, *Geophysical Research Letters*, *31*(19).

Rienecker, M. M., M. J. Suarez, R. Gelaro, R. Todling, J. Bacmeister, E. Liu, M. G. Bosilovich, S. D. Schubert, L. Takacs, G.-K. Kim, et al. (2011), MERRA: NASA's modern-era retrospective analysis for research and applications, *Journal of Climate*, *24*(14), 3624–3648.

Rodell, M., I. Velicogna, and J. S. Famiglietti (2009), Satellite-based estimates of groundwater depletion in india, *Nature*, *460*(7258), 999–1002.

Sen, P. K. (1968), Estimates of the regression coefficient based on Kendall's tau, *Journal of the American Statistical Association*, *63*(324), 1379–1389.

Shamsudduha, M., R. Taylor, and L. Longuevergne (2012), Monitoring groundwater storage changes in the highly seasonal humid tropics: Validation of GRACE measurements in the Bengal Basin, *Water Resources Research*, *48*(2).

Siebert, S., V. Henrich, K. Frenken, and J. Burke (2013), Update of the digital global map of irrigation areas to version 5, *Rheinische Friedrich-Wilhelms-Universität, Bonn, Germany and Food and Agriculture Organization of the United Nations, Rome, Italy*.

Soni, A., and T. H. Syed (2015), Diagnosing Land Water Storage Variations in Major Indian River Basins using GRACE observations, *Global and Planetary Change*, *133*, 263–271.

Tapley, B. D., S. Bettadpur, J. C. Ries, P. F. Thompson, and M. M. Watkins (2004), Grace measurements of mass variability in the earth system, *Science*, *305*(5683), 503–505.

Tiwari, V., J. Wahr, and S. Swenson (2009), Dwindling groundwater resources in northern India, from satellite gravity observations, *Geophysical Research Letters*, *36*(18).

Zaitchik, B. F., M. Rodell, and R. H. Reichle (2008), Assimilation of GRACE terrestrial water storage data into a land surface model: Results for the Mississippi River basin, *Journal of Hydrometeorology*, *9*(3), 535–548.

Zaveri, E., D. S. Grogan, K. Fisher-Vanden, S. Frolking, R. B. Lammers, D. H. Wrenn, A. Prusevich, and R. E. Nicholas (2016), Invisible water, visible impact: groundwater use and indian agriculture under climate change, *Environmental Research Letters*, *11*(8), 084,005.

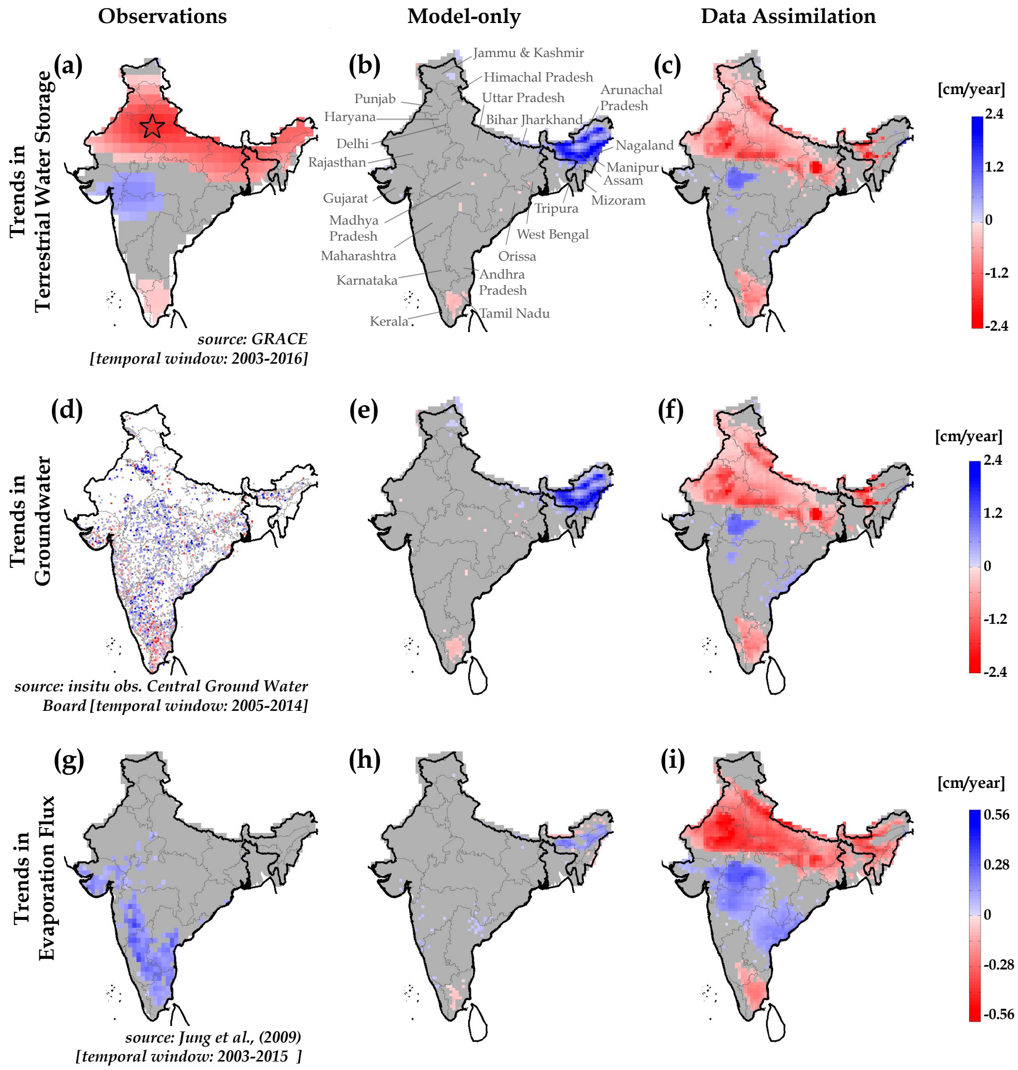
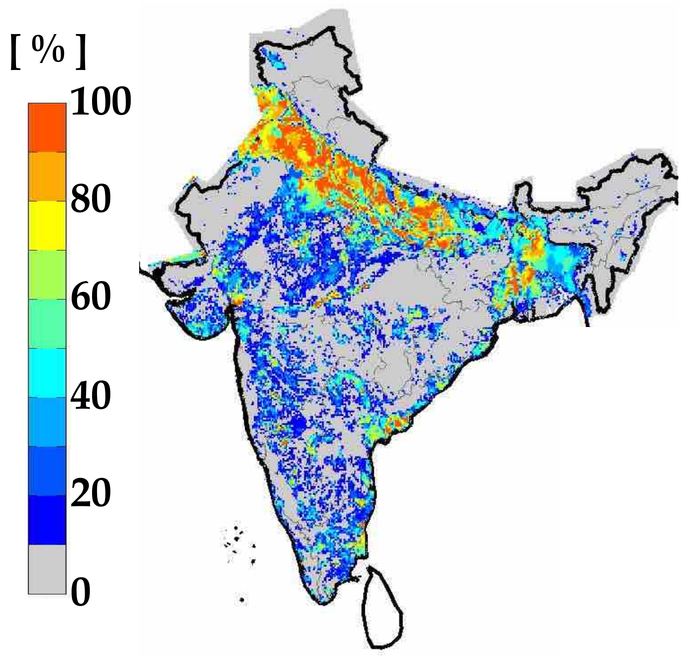


Figure 1. Trends in the (a,d,g) observed, (b,e,h) model-only, and (c,f,i) data assimilation estimates of (a,b,c) TWS, (d,e,f) groundwater, and (g,h,i) evapotranspiration rate. The “star" marker in (a) indicates the location of the time series shown in **Figure 3**. Grey colors indicate non-significant trends (p<0.05).



442 **Figure 2.** Percentage of land area equipped for irrigation, around the year 2005 [*Siebert et al.*, 2013].

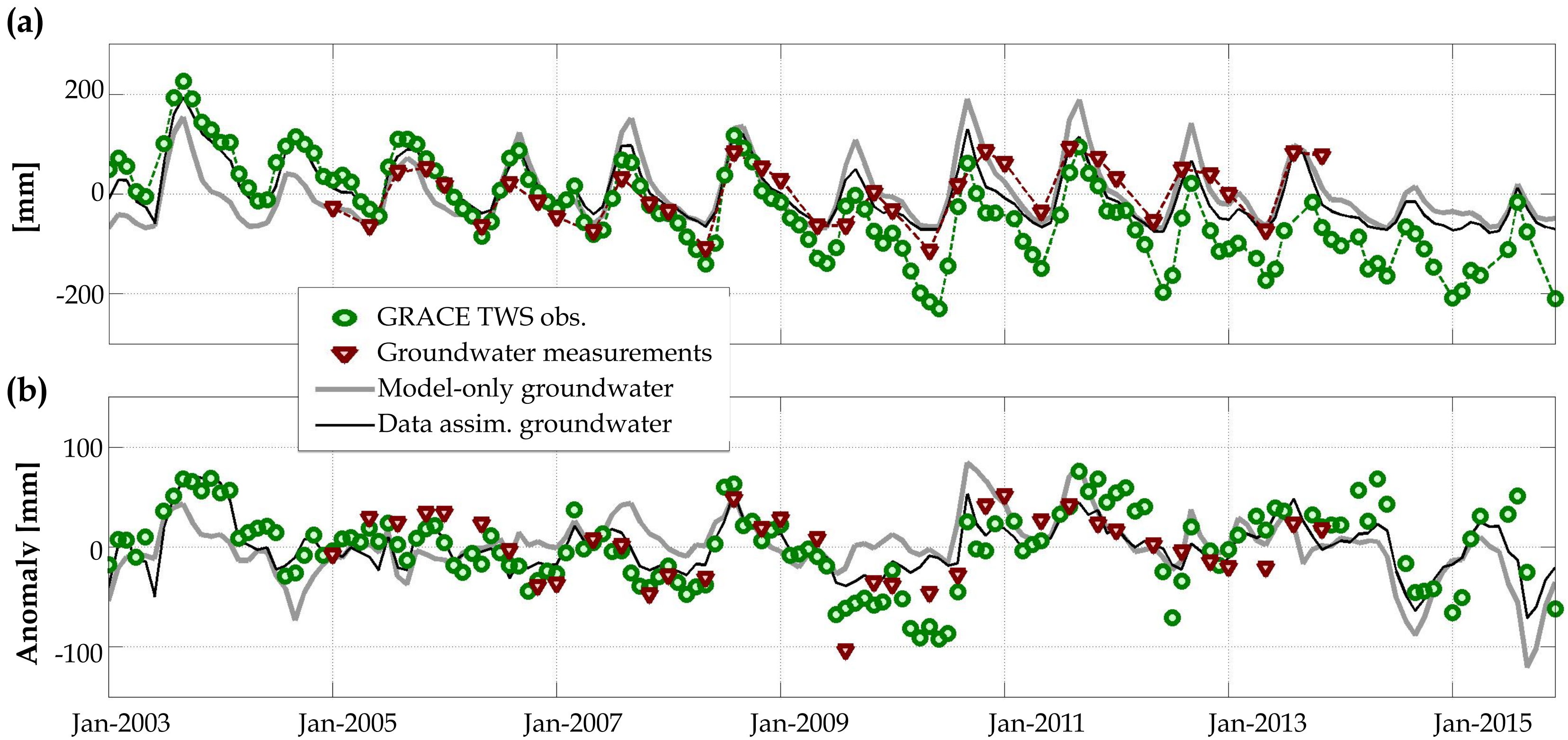


Figure 3. (a) (Green circles) GRACE TWS observations, (red triangles) in situ groundwater measurements, (thick grey line) model-only groundwater, and (black line) groundwater estimates from data assimilation for the location with the maximum TWS trend in GRACE observations (marked in **Figure 1a**). (b) As in (a) but for anomalies (with trends and the mean seasonal cycle removed). For this illustration, all data are aggregated from the 36 km model grid to the resolution of GRACE TWS observations.

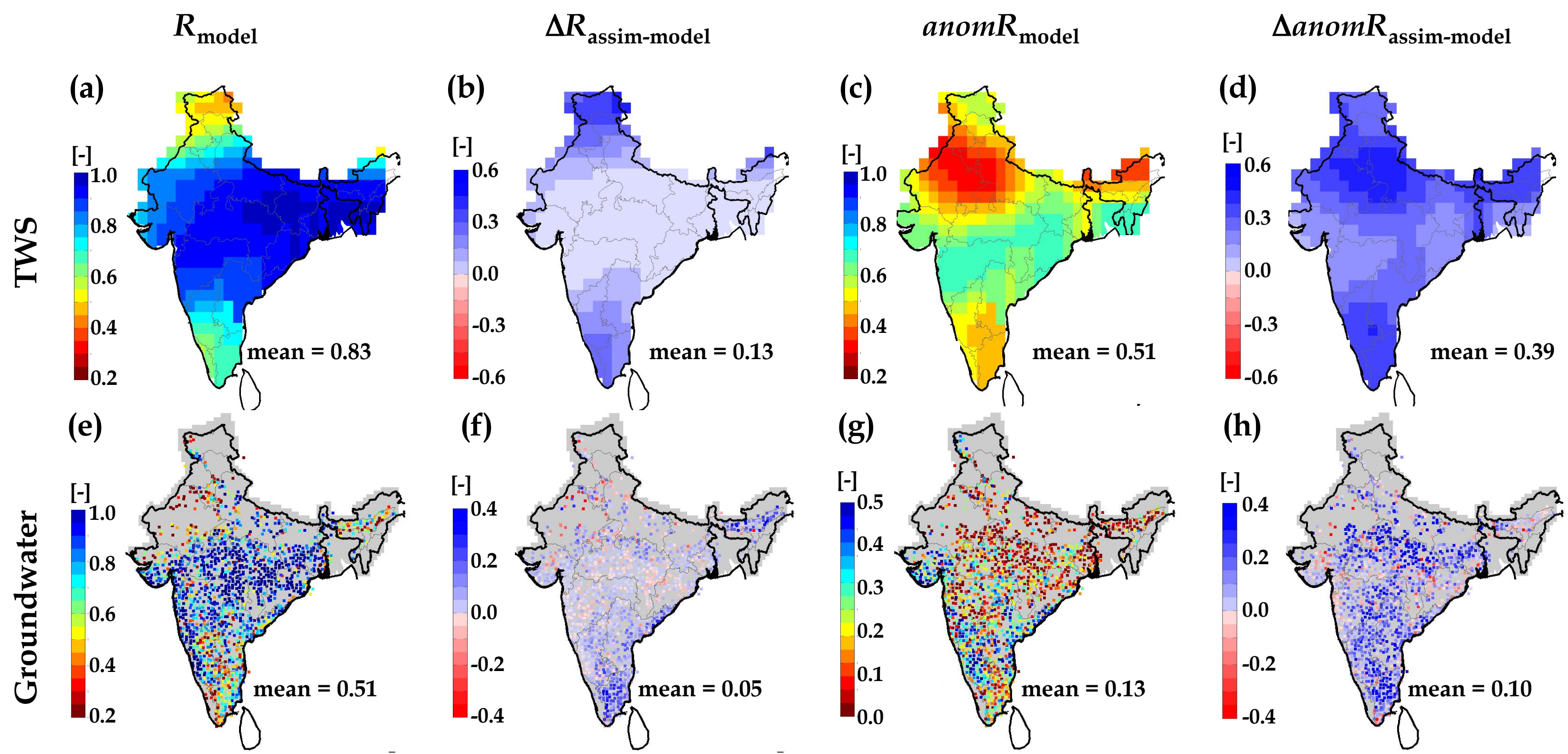


Figure 4. (a,e) Correlation (R) and (c,g) anomaly correlation ($anomR$) for model-only (a,c) TWS and (e,g) groundwater. Differences in (b,f) correlation (ΔR) and (d,h) anomaly correlation ($\Delta anomR$) between the assimilation and the model-only experiment for (b,d) TWS and (f,h) groundwater. Blue colors in skill difference plots (b,d,f,h) indicate that assimilation estimates are improved compared to model-only estimates, and red colors indicate that assimilation estimates are degraded. Numerical values provide area-average statistics (Section 2.5).

Supporting Information for

“Benefits and Pitfalls of GRACE Data Assimilation: a Case Study of Terrestrial Water Storage Depletion in India”

Manuela Girotto[1,2], Gabriëlle J. M. De Lannoy[3], Rolf H. Reichle[1], Matthew Rodell[4], Clara Draper[5,6], Soumendra N. Bhanja[7], Abhijit Mukherjee[7,8]

[1]Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, MD, USA.

[2]GESTAR, Universities Space Research Association, Columbia, MD 21044, USA.

[3]Department of Earth and Environmental Sciences, KU Leuven, Belgium.

[4]Hydrological Science Lab, NASA Goddard Space Flight Center, Greenbelt, MD, USA.

[5]Physical Sciences Division, NOAA/Earth System Research Laboratory, Boulder, Colorado.

[6]Cooperative Institute for Research in the Environmental Sciences, University of Colorado Boulder, Boulder, Colorado.

[7]Department of Geology and Geophysics, Indian Institute of Technology, Kharagpur, India.

[8]School of Environmental Science and Engineering, Indian Institute of Technology Kharagpur, West Bengala, India.

Contents

Figure S1.

Jan. 2003 - Dec. 2015 average of (a) monthly mean MERRA precipitation, (b) standard deviation in monthly observed (GRACE) TWS and (c) standard deviation in monthly model TWS. In the main paper, we refer to wet and dry areas where the mean precipitation is more or less than the average over India, respectively.

Table S1.

Ensemble perturbation parameters. Multiplicative (M) or Additive (A) perturbations are applied to precipitation (pcp), incoming solar radiation (sw), incoming longwave radiation (lw), catchment deficit (catdef), surface excess (srfexc), and snow water equivalent (swe). Spatial correlations are indicated as x, y_{corr} and temporal correlations as t_{corr}.

Corresponding author: Manuela Girotto, manuela.girotto@nasa.gov

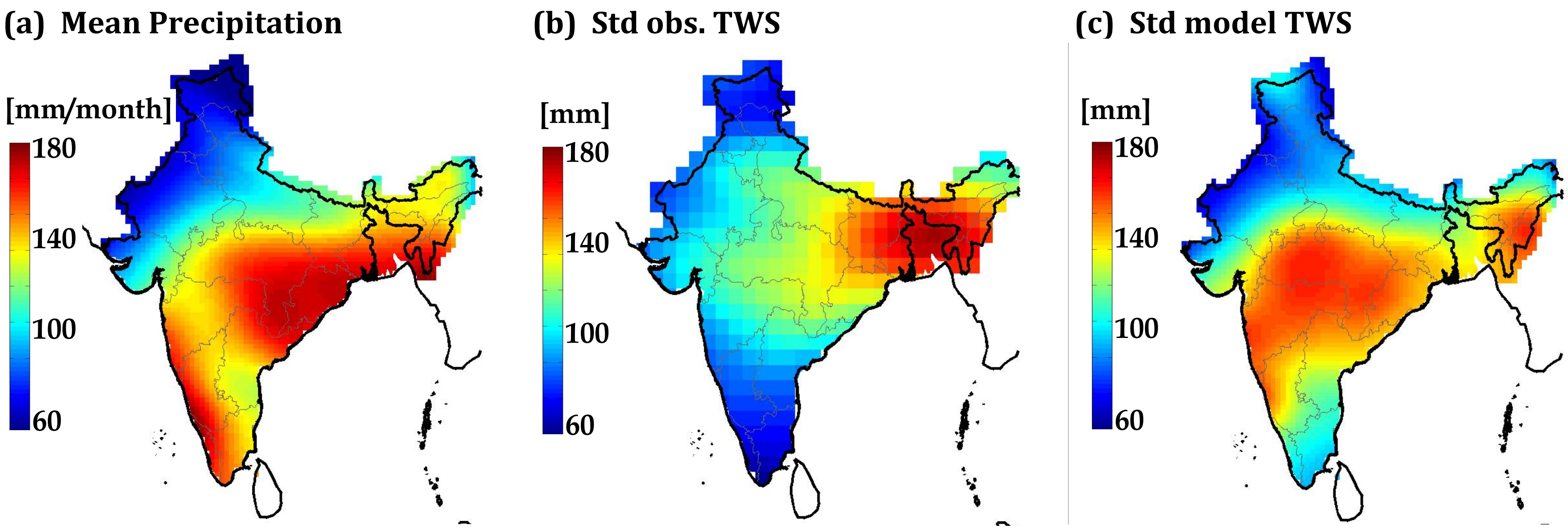


Figure 1. Jan. 2003 - Dec. 2015 average of (a) monthly mean MERRA precipitation, (b) standard deviation in monthly observed (GRACE) TWS and (c) standard deviation in monthly model TWS. In the main paper, we refer to wet and dry areas where the mean precipitation is more or less than the average over India, respectively.

Table 1. Ensemble perturbation parameters. Multiplicative (M) or Additive (A) perturbations are applied to precipitation (pcp), incoming solar radiation (sw), incoming longwave radiation (lw), catchment deficit (catdef), surface excess (srfexc), and snow water equivalent (swe). Spatial correlations are indicated as x, y_{corr} and temporal correlations as t_{corr}.

					cross-corr. with perturbations in		
	type	**standard deviation**	**x,y$_{corr}$**	**t$_{corr}$**	**pcp**	**sw**	**lw**
pcp	M	0.5	2°	3 days	n/a	-0.8	0.5
sw	M	0.3	2°	3 days	-0.8	n/a	-0.5
lw	A	20 Wm^{-2}	2°	3 days	0.5	-0.5	n/a
catdef	A	0.30 kg m^{-2} hr^{-1}	2°	1 days			
srfexc	A	0.06 kg m^{-2} hr^{-1}	2°	1 days			
swe	M	0.0012	2°	1 days			